



# UNIVERSIDAD DE LA RIOJA

## TRABAJO FIN DE ESTUDIOS

Título

Series Temporales. Aplicación al número de viajeros en La Rioja en establecimientos hoteleros 1999 - 2017.

Autor/es

GUILLERMO QUESADA VIGUERA

Director/es

Zenaida Hernández Martín y JUAN CARLOS FILLAT BALLESTEROS ,

Facultad

Facultad de Ciencias Empresariales

Titulación

Grado en Administración y Dirección de Empresas

Departamento

MATEMÁTICAS Y COMPUTACIÓN

Curso académico

2017-18



***Series Temporales. Aplicación al número de viajeros en La Rioja en establecimientos hoteleros 1999 - 2017.***, de GUILLERMO QUESADA VIGUERA (publicada por la Universidad de La Rioja) se difunde bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported. Permisos que vayan más allá de lo cubierto por esta licencia pueden solicitarse a los titulares del copyright.



FACULTAD DE CIENCIAS EMPRESARIALES

**TRABAJO FIN DE GRADO**

**GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS**

**Series Temporales. Aplicación al número de viajeros en  
La Rioja en establecimientos hoteleros 1999-2017.**

**Time Series. Application to the number of travelers in La  
Rioja at hotel establishments 1999-2017.**

Autor: D.Guillermo Quesada Viguera

Tutores: D<sup>a</sup>. Zenaida Hernández Martín

D. Juan Carlos Fillat Ballesteros

**CURSO ACADÉMICO 2017-2018**

# ÍNDICE

1.	INTRODUCCIÓN .....	4
2.	COMPONENTES DE LAS SERIES TEMPORALES .....	5
2.1.	Análisis de la Tendencia Secular .....	6
2.1.1.	Método de las medias móviles .....	7
2.1.2.	Método del ajuste mínimo cuadrático .....	9
2.2.	Análisis de la variación estacional .....	12
	Obtención de los Índices de Variación Estacional .....	12
2.3.	Desestacionalización de series de tiempo .....	13
3.	APLICACIÓN A UN CASO REAL EN LA ECONOMÍA RIOJANA: Número de viajeros en establecimientos hoteleros 1999-2017. ....	14
3.1.	Presentación de la base de datos .....	14
3.2.	Cálculo de la tendencia secular $T_t$ .....	16
3.2.1.	Cálculo de tendencia mediante el método de medias móviles .....	16
3.2.2.	Cálculo de tendencia mediante el método mínimo cuadrático.....	18
3.3.	Cálculo y análisis de la variación estacional.....	23
3.4.	Desestacionalización de la serie.....	25
3.5	Predicción del Número de Viajeros en La Rioja. Enero – Mayo 2018.....	26
4.	CONCLUSIONES .....	28
5.	BIBLIOGRAFÍA.....	30

## **RESUMEN**

El estudio se desarrollará en torno a la Teoría clásica de Series Temporales, el establecimiento de los componentes que explican el comportamiento de las mismas y la metodología para el aislamiento y estudio de cada uno de ellos, definidos como Tendencia, variaciones Estacionales, Cíclicas y Accidentales. Para ello, se expondrán los métodos de medias móviles y mínimo cuadrático, referente a la Tendencia, y por otro lado la obtención de los Índices de Variación Estacional y Desestacionalización, respecto al componente estacional.

Una vez analizada y expuesta dicha teoría comenzará la aplicación de la misma a un caso de la economía riojana, en concreto al número de viajeros en establecimientos hoteleros en La Rioja desde enero de 1999 hasta diciembre de 2017, el estudio recorrerá los componentes que explican el comportamiento de las series temporales, a la vez que relaciona los resultados obtenidos con los valores reales que componen la serie. Finalmente, y tras establecer un modelo representativo del comportamiento de la serie, dará lugar a la predicción de los primeros cinco meses de 2018, apreciando así la validez del estudio y del modelo estimado.

## **ABSTRACT**

The study will be developed around the classical Theory of Time Series, the establishment of the components that explain their behavior and the methodology for the isolation and study of each of them, defined as Trend, Seasonal, Cyclical and Accidental variations. For this purpose, the methods of moving averages and minimum square, referring to the Trend, and on the other hand, the estimation of the Seasonal Variation and Seasonalization Indices, regarding the seasonal component, will be presented.

Once this theory has been analysed and explained, it will begin to be applied to a case in the economy of La Rioja, specifically to the number of travelers in hotels in La Rioja from January 1999 to December 2017. The study will cover the components that explain the behaviour of the time series, while at the same time relating the results obtained to the real values that make up the series. Finally, after establishing a representative model for the behaviour of the series, it will lead to the prediction of the first five months of 2018, thus assessing the validity of the study and the estimated model.

# 1. INTRODUCCIÓN

El concepto de serie temporal se entiende dentro de la estadística descriptiva como toda colección de datos ordenados en el tiempo referidos a una determinada magnitud o variable, cuya referencia es una unidad temporal. Las series temporales son distribuciones estadísticas bidimensionales  $(t, Y)$  en las que una de las variables es el tiempo  $t$  y otra la variable de interés  $Y$ . Por tanto, el conjunto de datos que forman ambas variables vendrá presentado comúnmente en forma de tabla que contendrá la variable tiempo, representada por  $t_i$ , mientras que  $y_i$ , mostrará los valores de la variable de interés o estudio.

Teniendo en cuenta dicha definición es evidente que el análisis y tratamiento de series temporales es muy común en numerosas actividades humanas, y, por tanto, exigen el uso y estudio desde el punto de vista presentado. Algunos de los procesos más significativos que emplean dicho análisis son los económicos, demográficos o administrativos, entre otros.

Los principales objetivos que persigue el estudio clásico de series temporales son los siguientes:

- Obtención de un modelo que analice las variaciones pasadas y presentes del fenómeno estudiado
- Predicción de valores y situaciones futuras del mismo.

## 2. COMPONENTES DE LAS SERIES TEMPORALES

Según los métodos clásicos para el análisis y estudio de series temporales, se debe tener en cuenta la interacción de cuatro componentes, que determinarán los valores de la variable de interés en el transcurso temporal, y que son:

- **TENDENCIA SECULAR  $T_t$**  : consiste en la dirección principal que sigue la variable en el largo plazo. De este modo puede establecerse si el comportamiento de la misma obedece a un patrón, ya sea de crecimiento, decrecimiento, estabilidad o alternancia. La tendencia, por tanto, se determina sin tener en cuenta alteraciones puntuales en ciertos períodos de tiempo.

- **VARIACIONES ESTACIONALES  $S_t$**  : se trata de las oscilaciones descritas por la variable de interés ( $Y$ ) con cierta periodicidad. Generalmente suelen darse en períodos de tiempo (de un año o) inferiores a un año, tomando especial relevancia aquellas oscilaciones generadas de forma natural, a partir del simple transcurso estacional del año, equiparable a un fraccionamiento trimestral.

Por tanto, podemos establecer que las series más afectadas por la componente estacional serán aquellas que dependan en cierto grado de aspectos como el clima o hábitos humanos.

- **VARIACIONES CÍCLICAS  $C_t$**  : establece relación entre las fluctuaciones de la variable de interés ( $Y$ ) repetidas con periodicidad en intervalos temporales superiores al año.

Generalmente suelen presentarse en variables de carácter económico.

- **VARIACIONES ACCIDENTALES  $I_t$**  : recoge aquellas alteraciones bruscas e imprevistas de la trayectoria y que, como su nombre indica, son de carácter accidental. Además, debe añadirse que dichas alteraciones no son recogidas por el resto de componentes, y por tanto, pueden ser consideradas como deficiencias del modelo.

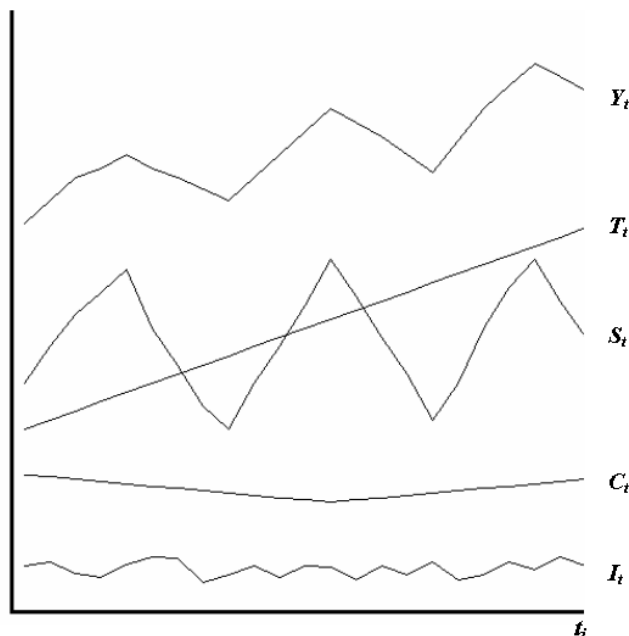
Una vez definidas las cuatro componentes, que determinan los valores de la variable de interés ( $Y_t$ ) en el horizonte temporal de forma individual y sin valorar con precisión la influencia de cada una de ellas dada su imposibilidad, se han establecido dos modelos de explicación, o hipótesis: la aditiva y la multiplicativa.

- **ADITIVA:**  $Y_t = T_t + S_t + C_t + I_t$

- **MULTIPLICATIVA:**  $Y_t = T_t \cdot S_t \cdot C_t \cdot I_t$

En el modelo multiplicativo, cabe destacar que las oscilaciones que experimenta la serie temporal pueden ser expresadas de forma porcentual respecto al valor de la Tendencia. Por tanto, las variaciones estacionales ( $S_t$ ) pueden ser expresadas como porcentajes tomando de referencia el 100%. Sin embargo, no es aplicable al modelo aditivo, ya que dichas fluctuaciones estarán expresadas en valor absoluto y en la unidad de medida de la propia variable de interés  $Y_t$ .

De este modo, la conjunción de las cuatro componentes expuestas puede interpretarse gráficamente en el siguiente gráfico:



Teniendo en cuenta lo expuesto anteriormente y debido a la complejidad del modelo, en el desarrollo del trabajo, se asumirá que las series temporales pueden ser descompuestas a través del modelo multiplicativo y que la componente de variaciones cíclicas  $C_t$  queda absorbida por la propia tendencia secular  $T_t$ . Presentándose el siguiente modelo:

$$Y_t = T_t \cdot S_t \cdot I_t$$

Además, la componente de estacionalidad se supondrá estable, es decir, que las fluctuaciones de carácter estacional se repiten con el mismo grado de variación todos los años.

## 2.1. Análisis de la Tendencia Secular

El análisis y cálculo de la tendencia secular  $T_t$  puede considerarse como el más influyente a largo plazo, permite la estimación de valores de la variable de interés  $Y_t$  para cualquier valor de  $t$ , estén o no presentes en la tabla de datos recogidos por la serie temporal.

Para su cálculo existen cuatro métodos básicos: Método gráfico, Método de las medias móviles, Método de las dos medias y Método del ajuste mínimo cuadrático. Sin embargo, en el desarrollo del trabajo únicamente será analizado el Método de las medias móviles y el Método del ajuste mínimo cuadrático, cabe destacar que este último es el método más preciso y exacto, y, por tanto, es particularmente útil para la estimación de valores futuros con cierto grado de certeza.



### 2.1.1. Método de las medias móviles

El objetivo principal será eliminar la influencia de cada una de las observaciones, para ello se promedia cada valor de la variable de interés  $Y_t$  con los valores contiguos a través de la media aritmética. De este modo se obtienen los valores que proporcionarán la tendencia secular  $T_t$ , y se generará así una distribución de menor dispersión que la original.

Para llevar a cabo dicho proceso deberá tenerse en cuenta el número de datos a promediar y se adoptará un número base para realizar el promedio, calificado como *tamaño de medias móviles*, siendo las más utilizadas aquellas medias de orden impar: 3, 5,... En el caso de utilizar un tamaño de medias móviles par, las medias extraídas inicialmente no estarán centradas, es decir, los valores que se obtienen se situarán entre un periodo y su contiguo, y por tanto será necesario volver a hacer el promedio de cada par de valores contiguos, para así obtener los valores centrados en cada período.

Por ejemplo, una serie temporal del tipo:

$t_i$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	$t_6$	$t_7$	$t_8$
$y_i$	$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$	$y_7$	$y_8$

Utilizando un *tamaño de medias móviles* igual a 3, se obtendrán las siguientes medias:

$$\bar{y}_2 = \frac{y_1 + y_2 + y_3}{3}$$

$$\bar{y}_5 = \frac{y_4 + y_5 + y_6}{3}$$

$$\bar{y}_3 = \frac{y_2 + y_3 + y_4}{3}$$

$$\bar{y}_6 = \frac{y_5 + y_6 + y_7}{3}$$

$$\bar{y}_4 = \frac{y_3 + y_4 + y_5}{3}$$

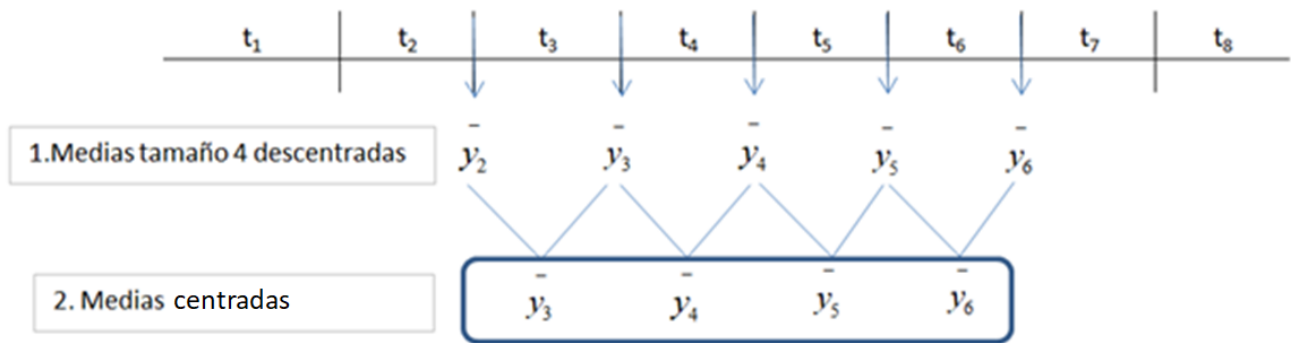
$$\bar{y}_7 = \frac{y_6 + y_7 + y_8}{3}$$

De este modo es evidente la inexistencia de las medias móviles para el valor del período 1 y 8, debido a que no existe valor ni anterior al 1, ni posterior al 8. Por tanto, la tendencia secular  $T_t$  de dicho ejemplo mediante el Método de medias móviles quedará como:

$\bar{t}_i$	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	$t_6$	$t_7$	$t_8$
$\bar{y}_i$	-	$\bar{y}_2$	$\bar{y}_3$	$\bar{y}_4$	$\bar{y}_5$	$\bar{y}_6$	$\bar{y}_7$	-

Si se tratara de un *tamaño de medias móviles* par, por ejemplo 4, deberíamos seguir el siguiente proceso:

- 1. Cálculo inicial de medias móviles de tamaño 4, valores descentrados.
- 2. Promedio de cada valor con su contiguo, y obtención del valor centrado.



Los valores obtenidos son un suavizado de la serie que nos permite observar la tendencia que sigue la variable estudiada en el largo plazo.

### 2.1.2. Método del ajuste mínimo cuadrático

El objetivo principal del método es la obtención de una función  $y_t^* = f(t)$  que relacione matemáticamente la variable de interés  $Y_t$  (variable dependiente) en función de la variable temporal  $t$  (variable independiente). Como he comentado anteriormente la consecución de dicha función se llevará a cabo mediante un ajuste mínimo cuadrático lineal, que dará lugar a la recta de regresión  $Y$  sobre  $t$ :

$$Y_t^* = a + bt$$

Por tanto, la tendencia que sigue la serie vendrá definida por los valores de la estimación de  $Y$ , es decir,  $y_t^*$ :

$$T_t = y_t^*$$

De este modo, los valores estimados mediante dicha recta se corresponderán con la tendencia lineal que sigue la serie de valores de la variable de interés  $Y_t$  a largo plazo, a través de

$$T_t = a + bt$$

expresión matemática que calificaremos como ecuación de tendencia, y cuyo análisis dependerá de la unidad temporal de la serie  $t$ , es decir:

- Si  $t$  está expresado en años la tendencia será anual.
- Si  $t$  está expresado en unidades temporales inferiores al año, meses, trimestres, semestres... (k-ésimos), la ecuación de tendencia generará la tendencia mensual, trimestral, semestral... (tendencia k-esimal).

Por tanto, a la hora de obtener la ecuación de tendencia deberemos tener en cuenta las dos posibilidades que se presentan en función de la unidad temporal de la serie.

#### 2.1.2.1 Obtención de la ecuación de la tendencia anual

Con el fin de obtener la ecuación de tendencia anual, la metodología general describe cuatro pasos a seguir:

Paso 1.- La medida temporal de la serie debe estar en años. Si no es así, y la serie viene definida en  $k$  períodos, se deberán agrupar en años y mediante el sumatorio de los valores de la variable de interés  $Y_t$  a lo largo de los  $k$  períodos de cada año (4 trimestres, 2 semestres, ...), pasando así la unidad temporal a anual.

Paso 2.- A continuación, y con el objetivo de simplificar la obtención de  $a$  y  $b$ , parámetros pertenecientes a la recta de regresión de  $Y$  sobre  $t$ , se llevará a cabo una transformación sobre la unidad temporal  $t$  expresada en años:

Se establecerá un año fijo, representado por  $t_0$ , y la variable  $t$  pasará a ser  $t'$ , del siguiente modo:  $t' = t - t_0$ . De este modo, se transforman en años enteros y consecutivos en torno al cero:

Si  $t_0$  es el primer año y los datos son de  $N$  años:  $t' = 0, 1, 2 \dots (N - 1)$

Paso 3.- Obtendremos la recta de regresión  $Y$  sobre  $t'$ :

$$Y_t^* = a + bt'$$

donde:

$$a = \bar{y} - b\bar{t'}$$

$$b = \frac{S_{t'Y}}{S_{t'}^2}$$

Paso 4.- La expresión que representará la ecuación de tendencia anual será:

$$T_{t'} = a + bt'$$

donde:  $t'$  está expresada en años,  $t' = 0, 1, 2 \dots (N - 1)$

cuyo origen es:  $t' = 0$  (parte central de  $t_0$ )

### Cambios de origen de la ecuación anual

Para llevar a cabo el proceso de cambio de origen en la ecuación de tendencia anual se deberá seguir el siguiente procedimiento:

Si la ecuación de tendencia anual  $T_{t'} = a + bt'$  cuenta con su origen en  $t = 0$ , es decir, el año  $t_0$  y se quiere realizar un traslado en el origen de  $h$  años a  $t = h$ , será en el año  $t_0 + h$  donde se lleve a cabo el cambio de origen sobre dicha ecuación:

$$t = t' + h$$

Por tanto, la ecuación de tendencia anual pasará a ser:

$$T_{t'} = a + b(t' + h) \quad \rightarrow \quad T_{t'} = (a + bh) + bt' \quad \text{con origen:}$$

$t' = 0$ , punto central del año  $(t_0 + h)$

### 2.1.2.2 Obtención de la ecuación de tendencia $k$ – esimal

Por tendencia  $k$  – esimal de la serie temporal, se entiende aquel estudio de la tendencia por unidades temporales inferiores al año, dichos períodos suelen coincidir con las divisiones naturales del año, es decir, semestres, trimestres, meses, ...  $k$ - ésimos generalmente. Por tanto, el valor que tome  $k$  supondrá el número de partes en las que se divide el año, por ejemplo, si  $k = 4$ , hablaremos de una división del año en 4 trimestres y por tanto la tendencia trimestral de la serie temporal.

Una vez definida la tendencia  $k$  – esimal, el proceso para su obtención partirá de la ecuación de tendencia anual y será el siguiente:

Paso 1.- Partiendo de la ecuación de tendencia anual  $T_t = a + bt$  la dividiremos por  $k$ :

$$T_t = \frac{a}{k} + \frac{b}{k}t$$

donde:  $t$  está expresada en años,  $t = 0, 1, 2 \dots (N - 1)$

$\frac{b}{k}$  muestra la tendencia  $k$  - esimal promedio del año  $t$ .

Paso 2.- Se dará el cambio de la unidad temporal: de años  $t$  a  $k$  – ésimos  $t'$

$$t' = kt$$

por tanto, la transformación consistirá en:  $t = \frac{t'}{k}$  dejando la ecuación como:

$$T_t^{(k)} = \frac{a}{k} + \frac{b}{k} \frac{t'}{k} \rightarrow T_t^{(k)} = \frac{a}{k} + \frac{b}{k^2} t' \quad \text{con origen: } t' = 0:$$

- si  $k$  es impar, en el  $k$  – ésimo central del año  $t = 0$

- si  $k$  es par, se situará entre los  $k$  – ésimos centrales

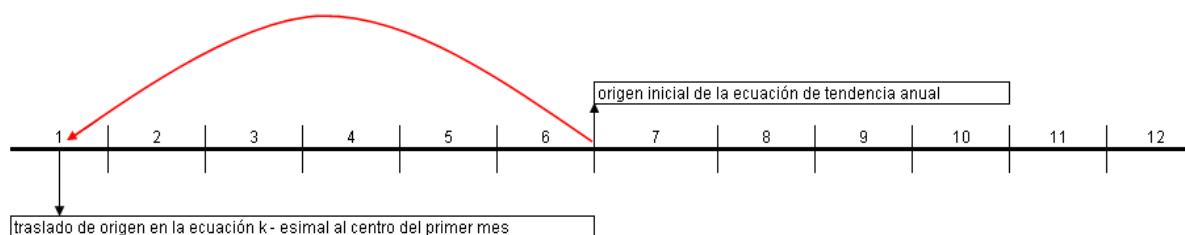
Con el fin evitar inconvenientes en el cálculo, deberemos fijar el origen de la ecuación de tendencia  $k$  – esimal en la parte central del primer  $k$  – ésimo.

#### Traslados del origen de la ecuación $k$ – esimal al centro del primer $k$ – ésimo

El proceso para llevar a cabo el traslado del origen de la ecuación de tendencia  $k$  – esimal al primer  $k$  – esimo del año de origen de la ecuación anual, deberemos aplicar el siguiente cambio, independientemente de si  $k$  es par o impar:

$$t'' = t' - \frac{k-1}{2}$$

Así por ejemplo si  $k = 12$ , el cambio será:  $t'' = t' - 5,5$



Una vez finalizado el estudio de la componente de Tendencia, continuaremos con el desarrollo y estudio de la componente Estacional.

## 2.2. Análisis de la variación estacional

La variación estacional en una serie temporal como ya hemos definido consiste en aquellas fluctuaciones u oscilaciones que provocan los “altibajos” en el gráfico de la serie. Con el análisis de la misma pretendemos llegar a la obtención de indicadores que respondan a situaciones del siguiente tipo:

La división del año en  $k$  – ésimos y los valores de la variable de interés para cada uno de ellos tienen un comportamiento de carácter regular y repetitivo a lo largo de los años de la serie temporal, es decir, estacionalidad estable. Por ejemplo, se produce una disminución de la variable de interés en el tercer trimestre y un posterior aumento en el cuarto trimestre, y se repite de un modo sistemático todos los años de la serie.

En un modelo multiplicativo  $Y_t = T_t \cdot S_t \cdot C_t \cdot I_t$ , el aspecto que tendrán los indicadores previamente mencionados será de índices y medirán la estacionalidad de la serie temporal en cada  $k$  – ésimo. A estos índices se les denominará como **Índices de Variación Estacional (IVE)** y se establecerá un índice para cada  $k$  en las que se divida el año. Cabe destacar, que dichos índices serán iguales para cada  $k$  – ésimo de cada año ya que adoptamos el supuesto de estacionalidad estable.

### Obtención de los Índices de Variación Estacional

La obtención de los IVE parte de aislar la componente de carácter estacional  $S_t$  de la serie temporal  $Y_t$ . Para llevar a cabo este proceso se dará por hecho que las variaciones cíclicas  $C_t$ , están absorbidas en la tendencia secular  $T_t$ .

Por tanto, el proceso para llevar a cabo el aislamiento de la componente estacional se conoce como el “método de la razón a la media móvil”, que consiste en lo siguiente:

Paso 1.- Se lleva a cabo la división de los valores  $k$  – esimales de la serie temporal  $Y_t$  por el valor obtenido en el análisis de la tendencia por el método de medias móviles  $T_t$  (suponiendo que esta contiene a  $C_t$ ):

$$\frac{Y_t}{T_t} = \frac{T_t C_t \cdot S_t \cdot I_t}{T_t C_t} = S_t \cdot I_t, \text{ valores que no contienen ni tendencia ni ciclo.}$$

Paso 2.- Por último, se promedian los valores anteriores, para cada  $k$  – ésimo en todos los años. De este modo, se considera que la componente accidental se reparte entre todo el período estudiado, quedando así aislada la componente de carácter estacional  $S_t$ .

Dichos promedios calculados, serán calificados como Movimientos Estacionales (ME) y a través de los mismos obtendremos los valores de los IVE, en porcentajes:

$$IVE(i) = \frac{ME(i)}{\sum_{i=1}^k ME(i)} \cdot k \cdot 100$$

#### *Aspecto y utilidad de los IVE*

Los Índices de Variación Estacional podrán presentarse en porcentaje o tanto por uno, por tanto, aquel IVE que presente un valor de 100 ó 1 supondrá que, en dicho  $k$  – ésimo, los valores de la serie no poseen un comportamiento de carácter estacional. Por el contrario, si el IVE muestra un valor superior o inferior a 100 en un  $k$  – ésimo determinado, se establecerá que en dicho  $k$  – ésimo, los valores de la serie son normalmente superiores o inferiores a los que deberían darse sin un efecto de carácter estacional.

Si el IVE = 100 → NO comportamiento estacional

Si el IVE  $\neq$  100 → SÍ comportamiento estacional

Por tanto, si  $T_t^{(k)}$  establece una predicción de la tendencia en el  $k$  – ésimo ( $i$ ), podremos corregir a través de la estacionalidad dicha predicción, multiplicándola por el IVE( $i$ ). Obteniendo así la predicción del valor real de la serie, al multiplicar la tendencia por la componente estacional correspondiente.

### **2.3. Desestacionalización de series de tiempo**

El proceso de desestacionalización de una serie temporal consiste en la eliminación de la componente estacional de los datos originales de dicha serie.

El método para llevarlo a cabo parte de la obtención de los  $k$  IVE, dividiendo los datos originales de la serie  $Y_t$  por los IVE ( $i$ ) de cada  $k$  – ésimo (en tanto por uno).

$$D_t = \frac{Y_t}{S_t} = \frac{Y_t}{IVE(i)}$$

Los valores corresponderán con la serie de datos desestacionalizada  $D_t$ , y consistirán en aquellos valores que se hubieran observado ante la ausencia de influencia estacional, es decir, el comportamiento que hubiera tenido la variable de interés ante la ausencia de “estaciones”.

### **3. APLICACIÓN A UN CASO REAL EN LA ECONOMÍA RIOJANA: Número de viajeros en establecimientos hoteleros 1999-2017.**

A continuación, llevaré a cabo una aplicación del análisis de series temporales, previamente detallado, sobre el número de Viajeros en establecimientos hoteleros en La Rioja desde enero de 1999 hasta diciembre de 2017. La elección de dicha variable recae en la propia importancia que representa el turismo en cualquier región o país como una de las actividades de carácter económico y cultural de mayor peso. En concreto, el análisis del número de viajeros en La Rioja durante las dos últimas décadas demuestra el crecimiento prolongado y sostenido del sector, y la toma de importancia continuada dentro de las limitaciones lógicas de la que es la región más pequeña de nuestro país. Pese a las limitaciones propias de una Comunidad Autónoma de interior La Rioja otorga y enfatiza la importancia del sector turístico dentro de su economía, cerrando el pasado 2017 con la mayor cifra histórica de turistas en nuestra región, 833.934, y mejorando así la cifra de 2016 en un 5,8%, asentando la tendencia de mejora y crecimiento tras el fuerte impacto de la crisis económica sufrida en nuestro país a finales de la pasada década.

El análisis de tipo estadístico de la serie temporal seleccionada seguirá las pautas de la teoría clásica de series temporales anteriormente expuesta, es decir, comenzará con la presentación de los datos a analizar, análisis de la Tendencia Secular mediante los métodos de medias móviles y mínimo cuadrático, análisis de la variación estacional y la desestacionalización de la misma. El objetivo principal será el de analizar y estudiar las variaciones pasadas y presentes de la variable indicada y el establecimiento de un modelo que permita la predicción y análisis de situaciones futuras.

#### **3.1. Presentación de la base de datos**

Los datos a analizar han sido recogidos de la base de datos del Instituto Nacional de Estadística, más concretamente pertenecen a los datos recogidos en la Encuesta de Ocupación Hotelera (EOH), dicha encuesta nos permite medir la evolución de la oferta y ocupación en alojamientos hoteleros mediante el análisis de numerosas variables como pernoctaciones, viajeros, nivel de ocupación, número de establecimientos, etc. Concretamente los datos corresponden con el número de viajeros mensuales en establecimientos hoteleros en La Rioja de 1999 a 2017.

Por un lado, la variable seleccionada ha sido la de *viajeros*, definida según el INE como: “*Todas aquellas personas que realizan una o más pernoctaciones seguidas en el mismo alojamiento*”. La elección de la misma es meramente por la mayor validez de los datos, ya que únicamente tiene en cuenta el número de personas que pernoctan en un establecimiento hotelero, y no el número de pernoctaciones que llevan a cabo cada una de ellas, como refleja la variable *pernoctaciones*.

La otra cuestión a justificar es que únicamente se presentan los viajeros que hayan pernoctado en *Establecimientos Hoteleros*, definida por el INE del siguiente modo: *Son todos los establecimientos hoteleros inscritos como tales en el correspondiente registro de las Consejerías de Turismo de cada Comunidad Autónoma. Son establecimientos hoteleros aquellos establecimientos que prestan*



*servicios de alojamiento colectivo mediante precio con o sin otros servicios complementarios (hotel, hotel-apartamento o apartahotel, motel, hostel, pensión...).* La elección de dicha tipología de establecimiento también recae en la mayor fiabilidad y validez de los datos recogidos en comparación con otros modelos de establecimiento como pueden ser apartamentos turísticos o albergues.

## Base de datos:

### Encuesta de ocupación hotelera

#### Viajeros en Establecimientos hoteleros en La Rioja 1999 - 2017

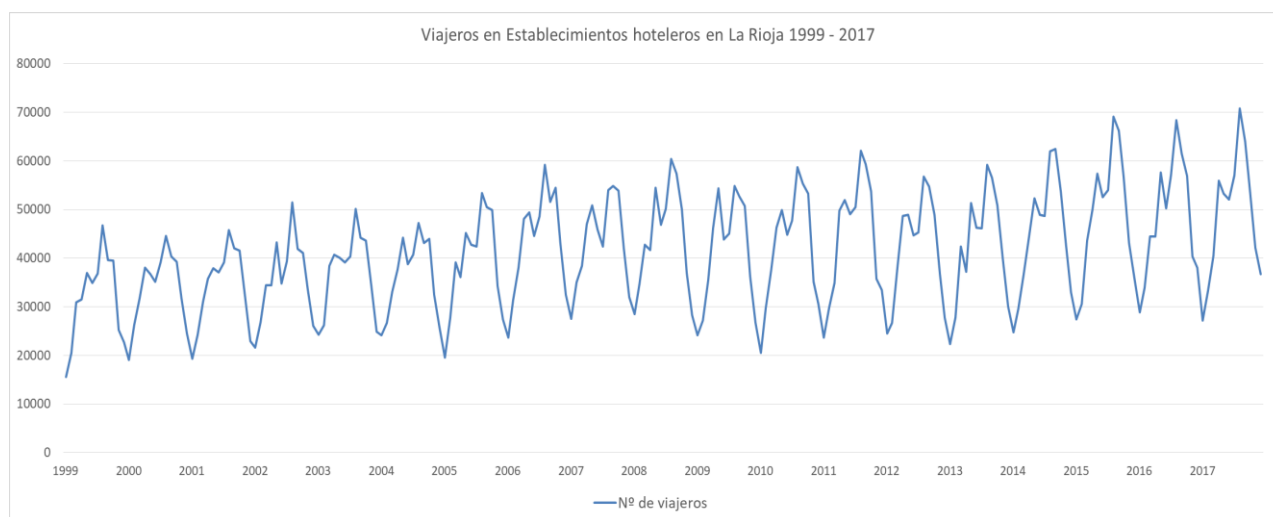
Unidades: Viajeros

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
M01- Enero	15.596	19.028	19.363	21.644	24.263	24.210	19.592	23.663	27.496	28.558	24.102	20.479	23.644	24.573	22.287	24.778	27.368	28.820	27.213
M02 - Febrero	20.480	26.313	24.284	26.801	26.211	26.708	27.779	31.444	35.057	34.634	27.174	29.730	29.777	26.636	27.819	29.553	30.611	33.921	33.389
M03 - Marzo	30.932	31.792	31.049	34.394	38.393	33.115	39.146	38.098	38.369	42.753	35.534	37.483	34.951	38.066	42.451	36.699	43.654	44.410	40.437
M04 - Abril	31.578	38.013	35.706	34.453	40.697	37.635	36.062	48.100	46.988	41.671	46.145	46.262	49.715	48.701	37.156	44.499	50.042	44.481	55.955
M05 - Mayo	37.023	36.732	37.893	43.287	40.076	44.211	45.179	49.466	50.835	54.531	54.304	49.914	51.972	48.922	51.331	52.261	57.351	57.634	53.247
M06 - Junio	34.866	35.153	37.133	34.848	39.139	38.809	42.708	44.549	45.938	46.824	43.871	44.863	49.002	44.647	46.286	48.962	52.532	50.307	52.062
M07 - Julio	36.880	39.157	39.135	39.357	40.362	40.766	42.374	48.606	42.428	50.102	45.094	47.710	50.540	45.345	46.107	48.655	53.959	57.005	56.988
M08 - Agosto	46.794	44.619	45.754	51.483	50.080	47.202	53.444	59.248	54.054	60.356	54.840	58.722	62.091	56.730	59.150	62.011	69.052	68.321	70.733
M09 - Septiembre	39.672	40.328	42.022	41.968	44.160	43.125	50.456	51.639	54.824	57.344	52.675	55.316	59.299	54.740	56.596	62.514	66.147	61.522	63.999
M10 - Octubre	39.529	39.200	41.607	41.101	43.559	43.932	49.936	54.463	53.892	50.012	50.688	53.230	53.753	48.984	50.890	53.786	56.654	56.952	53.860
M11 - Noviembre	25.298	31.616	32.264	33.471	34.401	32.455	34.276	42.935	41.893	37.000	36.299	35.191	35.766	36.871	40.039	42.445	43.149	40.323	42.201
M12 - Diciembre	22.744	24.516	23.000	26.119	24.908	25.778	27.453	32.478	32.043	28.216	26.647	30.522	33.511	27.739	30.096	32.929	35.933	38.050	36.688

Fuente:

Instituto Nacional de Estadística

- Gráficamente:



### 3.2. Cálculo de la tendencia secular $T_t$

Siguiendo el análisis de series temporales de la teoría clásica, analizaré en primer lugar la Tendencia Secular de la serie *Viajeros en establecimientos hoteleros en La Rioja 1999 – 2017*, cabe destacar que el estudio será mediante los dos métodos previamente expuestos, es decir, mediante *el método de medias móviles* y por *el método de ajuste mínimo cuadrático*. Previamente definía la Tendencia Secular como: la dirección principal que sigue la variable en el largo plazo. De este modo puede establecerse si el comportamiento de la misma obedece a un patrón, ya sea de crecimiento, decrecimiento, estabilidad o alternancia. La tendencia, por tanto, se determina sin tener en cuenta alteraciones puntuales en ciertos períodos de tiempo.

#### 3.2.1. Cálculo de tendencia mediante el método de medias móviles

El método de medias móviles consiste en establecer el promedio de cada valor con los valores contiguos del mismo mediante la media aritmética, generando una distribución de la variable con menos dispersión. Para llevarlo a cabo, se establece el *tamaño de medias móviles*, en este caso aplicaré un tamaño de 12, es decir, un año (12 meses). Al tratarse de un número par las medias resultantes serán valores no centrados, por lo que será necesario un segundo promedio de cada valor obtenido y su anterior:

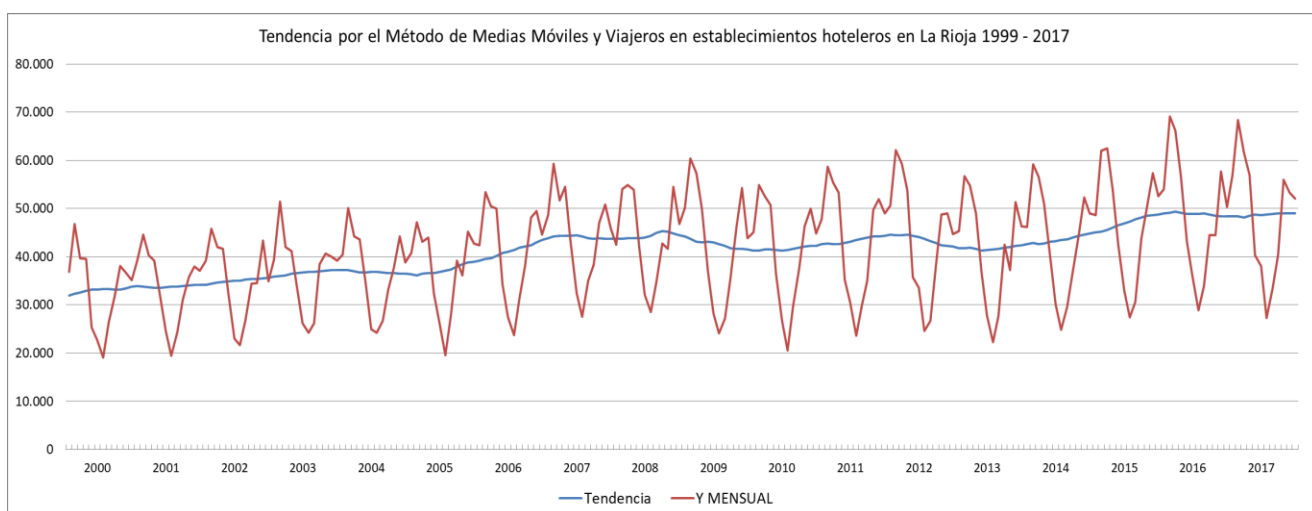
- 1. Cálculo de medias móviles descentradas:

MEDIAS MOVILES TAMAÑO 12 DESCENTRADAS												
	0,5	1,5	2,5	3,5	4,5	5,5	6,5	7,5	8,5	9,5	10,5	11,5
1999							31783	32069	32555	32626	33163	33138
2000	33162	33352	33171	33226	33198	33725	33872	33900	33731	33669	33477	33574
2001	33739	33737	33831	33973	34173	34227	34101	34291	34501	34779	34675	35125
2002	34934	34953	35430	35426	35383	35484	35744	35962	35913	36246	36767	36499
2003	36857	36940	36823	37006	37211	37288	37187	37183	37224	36785	36529	36874
2004	36847	36880	36640	36554	36585	36423	36496	36111	36200	36703	36571	36652
2005	36977	37111	37631	38242	38742	38894	39034	39373	39678	39591	40594	40952
2006	41105	41624	42108	42207	42584	43305	43724	44044	44345	44367	44275	44389
2007	44504	43990	43557	43822	43775	43688	43651	43740	43705	44070	43627	43935
2008	44009	44648	45173	45383	45060	44652	44333	43962	43340	42739	43112	43093
2009	42847	42429	41970	41581	41637	41579	41448	41146	41359	41521	41531	41165
2010	41248	41466	41789	42009	42221	42129	42452	42716	42720	42509	42796	42968
2011	43313	43549	43829	44161	44205	44253	44502	44579	44317	44577	44493	44238
2012	43875	43443	42996	42616	42218	42311	41830	41639	41738	42103	41141	41342
2013	41478	41542	41743	41898	42057	42321	42517	42725	42869	42390	43002	43080
2014	43303	43515	43753	44246	44488	44688	44924	45140	45228	45808	46270	46694
2015	46992	47434	48020	48323	48562	48621	48871	48992	49268	49331	48867	48891
2016	48706	48959	48899	48513	48538	48302	48479	48345	48301	47970	48926	48560
2017	48706	48705	48906	49112	48855	49011	48898					

- 2. Promedio de cada valor con su contiguo para la obtención de los valores centrados (parte central del mes):

MEDIAS MOVILES CENTRADAS TAMAÑO 12												
	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
1999							31926	32312	32591	32895	33151	33150
2000	33257	33261	33198	33212	33461	33798	33886	33816	33700	33573	33525	33656
2001	33738	33784	33902	34073	34200	34164	34196	34396	34640	34727	34900	35029
2002	34943	35191	35428	35404	35434	35614	35853	35938	36080	36506	36633	36678
2003	36898	36882	36915	37108	37250	37238	37185	37204	37005	36657	36702	36860
2004	36863	36760	36597	36570	36504	36459	36303	36155	36451	36637	36612	36815
2005	37044	37371	37937	38492	38818	38964	39203	39526	39635	40093	40773	41028
2006	41365	41866	42157	42395	42945	43515	43884	44194	44356	44321	44332	44446
2007	44247	43773	43689	43798	43731	43670	43696	43722	43887	43848	43781	43972
2008	44329	44911	45278	45222	44856	44493	44148	43651	43040	42925	43102	42970
2009	42638	42200	41775	41609	41608	41513	41297	41252	41440	41526	41348	41207
2010	41357	41628	41899	42115	42175	42290	42584	42718	42614	42652	42882	43140
2011	43431	43689	43995	44183	44229	44377	44540	44448	44447	44535	44365	44057
2012	43659	43219	42806	42417	42264	42070	41734	41688	41920	41622	41241	41410
2013	41510	41643	41821	41978	42189	42419	42621	42797	42630	42696	43041	43191
2014	43409	43634	44000	44367	44588	44806	45032	45184	45518	46039	46482	46843
2015	47213	47727	48172	48443	48591	48746	48932	49130	49299	49099	48879	48798
2016	48833	48929	48706	48526	48420	48391	48412	48323	48135	48448	48743	48633
2017	48706	48805	49009	48984	48933	48954						

- Gráficamente:



### 3.2.2. Cálculo de tendencia mediante el método mínimo cuadrático

El cálculo de la tendencia a través de esta herramienta tiene como objetivo alcanzar una función  $y_t^* = f(t)$  que contenga y relacione la variable de interés  $Y_t$ , en este caso Número de Viajeros (variable dependiente), en función de la variable temporal  $t$  (variable independiente). De este modo, mediante el ajuste mínimo cuadrático lineal se obtendrá la recta de regresión  $Y$  sobre  $t$ , es decir,  $Y_t^* = a + bt$ , siendo la estimación de  $Y_t$ , los valores correspondientes a la Tendencia,  $T_t = y_t^*$ . Por tanto, la obtención de la Tendencia vendrá definida por  $T_t = a + bt$ .

#### 3.2.2.1 Obtención de la ecuación de tendencia anual

La base de datos a analizar corresponde con una serie mensual, por tanto, el objetivo a alcanzar será el cálculo de la tendencia mensual de la serie. Pese a ello, será necesario en primer lugar hallar la tendencia anual para realizar el posterior cambio y obtención de la tendencia mensual.

- 1. Paso de la unidad temporal mensual a anual
- 2. Establecimiento de 2008 como origen, la decisión de tomar el valor central como base recae simplemente en la obtención de números más simples en los cálculos, quedará representado por  $t_0$ , y la variable  $t$  pasará a ser  $t'$ , del siguiente modo:  $t' = t - t_0$ . De este modo, se transforman los años de la serie en años enteros y consecutivos en torno al cero:

AÑO	Nº Anual de Viajeros
1999	381.392
2000	406.467
2001	409.210
2002	428.926
2003	446.249
2004	437.946
2005	468.405
2006	524.689
2007	523.817
2008	532.001
2009	497.373
2010	509.422
2011	534.021
2012	501.954
2013	510.208
2014	539.092
2015	586.452
2016	581.746
2017	586.772

AÑO	año $t' = t - 2008$
1999	-9
2000	-8
2001	-7
2002	-6
2003	-5
2004	-4
2005	-3
2006	-2
2007	-1
2008	0
2009	1
2010	2
2011	3
2012	4
2013	5
2014	6
2015	7
2016	8
2017	9

- 3. Cálculos para la obtención de la recta de regresión:  $Y_t^* = a + bt'$

AÑO	año t'	Nº Anual de Viajeros (Y)	Y <sup>2</sup>	t' <sup>2</sup>	Y · t'
1999	-9	381.392	1,4546E+11	81	-3.432.528,00
2000	-8	406.467	1,65215E+11	64	-3.251.736,00
2001	-7	409.210	1,67453E+11	49	-2.864.470,00
2002	-6	428.926	1,83978E+11	36	-2.573.556,00
2003	-5	446.249	1,99138E+11	25	-2.231.245,00
2004	-4	437.946	1,91797E+11	16	-1.751.784,00
2005	-3	468.405	2,19403E+11	9	-1.405.215,00
2006	-2	524.689	2,75299E+11	4	-1.049.378,00
2007	-1	523.817	2,74384E+11	1	-523.817,00
2008	0	532.001	2,83025E+11	0	0,00
2009	1	497.373	2,4738E+11	1	497.373,00
2010	2	509.422	2,59511E+11	4	1.018.844,00
2011	3	534.021	2,85178E+11	9	1.602.063,00
2012	4	501.954	2,51958E+11	16	2.007.816,00
2013	5	510.208	2,60312E+11	25	2.551.040,00
2014	6	539.092	2,9062E+11	36	3.234.552,00
2015	7	586.452	3,43926E+11	49	4.105.164,00
2016	8	581.746	3,38428E+11	64	4.653.968,00
2017	9	586.772	3,44301E+11	81	5.280.948,00
TOTAL		9.406.142	4,72677E+12	570	5.868.039,00
MEDIAS		495.060	2,48777E+11	30	308.844,16

$$Y_t^* = a + bt' \quad \left\{ \begin{array}{l} \sum Y_i = a \cdot N \\ \sum t_i' Y_i = b \cdot \sum t_i'^2 \end{array} \right.$$

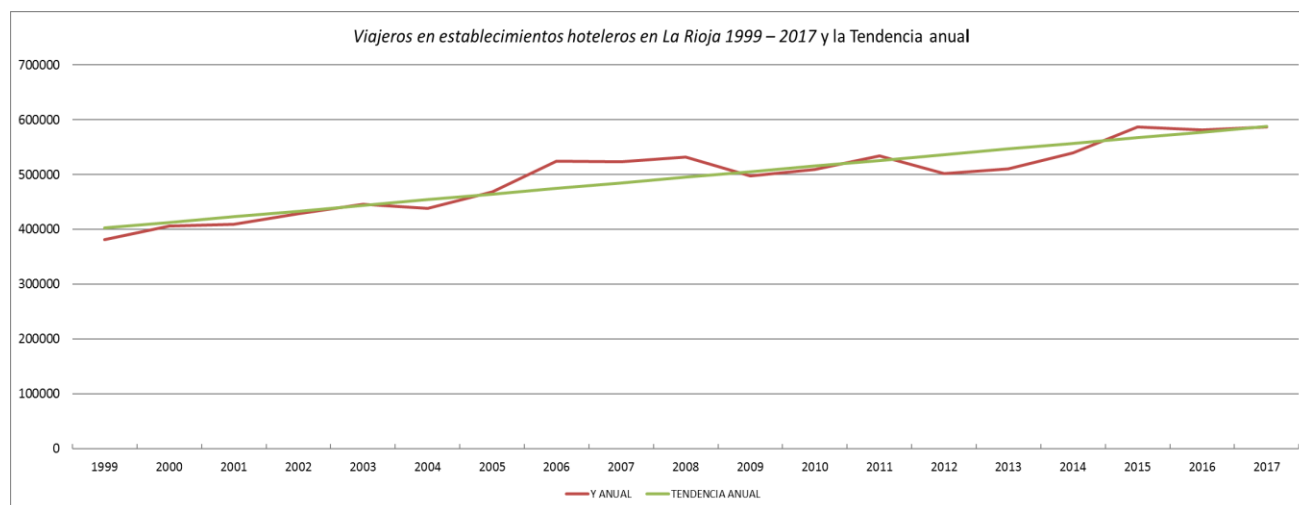
Por tanto, la tendencia anual en base a 2008 que sigue la serie vendrá definida por los valores de la estimación de Y, es decir,  $Y_t^*$ :

$$T_t = a + bt \quad \rightarrow \quad T_t = 495060,105 + 10294,805 \cdot t'$$

- 4. Obtención de valores de la Tendencia anual a través de la recta de regresión estimada:

AÑO	TENDENCIA ANUAL
1999	402406,9
2000	412701,7
2001	422996,5
2002	433291,3
2003	443586,1
2004	453880,9
2005	464175,7
2006	474470,5
2007	484765,3
2008	495060,1
2009	505354,9
2010	515649,7
2011	525944,5
2012	536239,3
2013	546534,1
2014	556828,9
2015	567123,7
2016	577418,5
2017	587713,4

- 5. Gráfico comparativo de los valores reales de la serie anual *Viajeros en establecimientos hoteleros en La Rioja 1999 – 2017* y la Tendencia anual estimada:



#### *Obtención del coeficiente de determinación $R^2$ para la ecuación de tendencia anual*

La obtención del coeficiente de determinación nos permite a través de su cálculo determinar la calidad y validez del modelo establecido, y la proporción en que varían los resultados explicada por dicho modelo. Dicho estadístico es comúnmente utilizado en la predicción de resultados, así como en la comprobación de hipótesis, como la que será llevada a cabo más adelante.

- Cálculo:

$$R^2 = \frac{\sum_{t=1}^T (\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

- Interpretación del Coeficiente de determinación:  **$R^2 = 0.861$**

El coeficiente de determinación únicamente puede tomar valores entre cero y uno, en nuestro caso 0.861, es un valor bastante próximo a uno y por tanto denota un grado de calidad del modelo elevado. En concreto, hace referencia al grado de relación entre ambas variables de nuestro modelo, tiempo y número de viajeros, de tal modo, que en nuestro caso podemos asegurar que el número de viajeros puede ser explicado en un 86,1% por la variable temporal. Del mismo modo podremos asegurar que la predicción que será llevada a cabo más adelante será representativa y de una fiabilidad considerable.

### 3.2.2.2 Obtención de la ecuación de tendencia mensual

Una vez definida la tendencia anual  $T_t = 49506Q105 + 10294805 \cdot t'$ , el proceso para la obtención de la tendencia mensual  $T_t^{(12)} = \frac{a}{12} + \frac{b}{12} \frac{t'}{12}$  seguirá el siguiente proceso:

- 1. Estimación de la recta de regresión de tendencia mensual:

$$T_t^{(12)} = \frac{49506Q105}{12} + \frac{10294805}{12^2} \cdot t''$$

- 2. Transformación de los meses de la serie en números enteros y consecutivos en torno al cero:

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
1999	-108	-107	-106	-105	-104	-103	-102	-101	-100	-99	-98	-97
2000	-96	-95	-94	-93	-92	-91	-90	-89	-88	-87	-86	-85
2001	-84	-83	-82	-81	-80	-79	-78	-77	-76	-75	-74	-73
2002	-72	-71	-70	-69	-68	-67	-66	-65	-64	-63	-62	-61
2003	-60	-59	-58	-57	-56	-55	-54	-53	-52	-51	-50	-49
2004	-48	-47	-46	-45	-44	-43	-42	-41	-40	-39	-38	-37
2005	-36	-35	-34	-33	-32	-31	-30	-29	-28	-27	-26	-25
2006	-24	-23	-22	-21	-20	-19	-18	-17	-16	-15	-14	-13
2007	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
2008	0	1	2	3	4	5	6	7	8	9	10	11
2009	12	13	14	15	16	17	18	19	20	21	22	23
2010	24	25	26	27	28	29	30	31	32	33	34	35
2011	36	37	38	39	40	41	42	43	44	45	46	47
2012	48	49	50	51	52	53	54	55	56	57	58	59
2013	60	61	62	63	64	65	66	67	68	69	70	71
2014	72	73	74	75	76	77	78	79	80	81	82	83
2015	84	85	86	87	88	89	90	91	92	93	94	95
2016	96	97	98	99	100	101	102	103	104	105	106	107
2017	108	109	110	111	112	113	114	115	116	117	118	119

- 3. Traslado de origen al centro del primer mes, enero de 2008:  $t'' = t' - \frac{k-1}{2}$  ;  $t'' = t' - \frac{12-1}{2}$

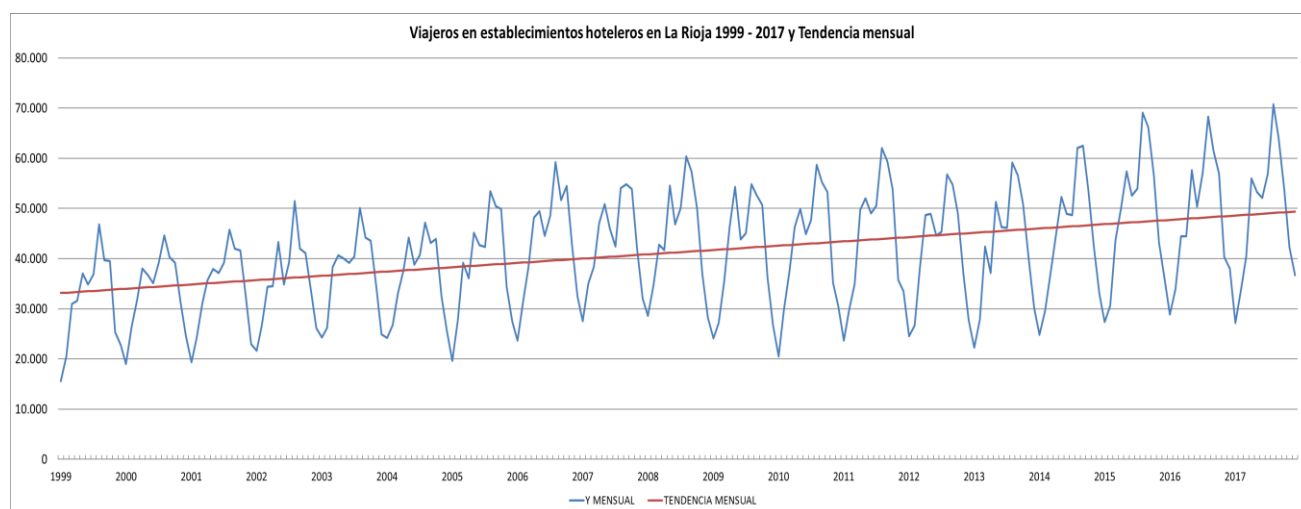
	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
1999	-113,5	-112,5	-111,5	-110,5	-109,5	-108,5	-107,5	-106,5	-105,5	-104,5	-103,5	-102,5
2000	-101,5	-100,5	-99,5	-98,5	-97,5	-96,5	-95,5	-94,5	-93,5	-92,5	-91,5	-90,5
2001	-89,5	-88,5	-87,5	-86,5	-85,5	-84,5	-83,5	-82,5	-81,5	-80,5	-79,5	-78,5
2002	-77,5	-76,5	-75,5	-74,5	-73,5	-72,5	-71,5	-70,5	-69,5	-68,5	-67,5	-66,5
2003	-65,5	-64,5	-63,5	-62,5	-61,5	-60,5	-59,5	-58,5	-57,5	-56,5	-55,5	-54,5
2004	-53,5	-52,5	-51,5	-50,5	-49,5	-48,5	-47,5	-46,5	-45,5	-44,5	-43,5	-42,5
2005	-41,5	-40,5	-39,5	-38,5	-37,5	-36,5	-35,5	-34,5	-33,5	-32,5	-31,5	-30,5
2006	-29,5	-28,5	-27,5	-26,5	-25,5	-24,5	-23,5	-22,5	-21,5	-20,5	-19,5	-18,5
2007	-17,5	-16,5	-15,5	-14,5	-13,5	-12,5	-11,5	-10,5	-9,5	-8,5	-7,5	-6,5
2008	-5,5	-4,5	-3,5	-2,5	-1,5	-0,5	0,5	1,5	2,5	3,5	4,5	5,5
2009	6,5	7,5	8,5	9,5	10,5	11,5	12,5	13,5	14,5	15,5	16,5	17,5
2010	18,5	19,5	20,5	21,5	22,5	23,5	24,5	25,5	26,5	27,5	28,5	29,5
2011	30,5	31,5	32,5	33,5	34,5	35,5	36,5	37,5	38,5	39,5	40,5	41,5
2012	42,5	43,5	44,5	45,5	46,5	47,5	48,5	49,5	50,5	51,5	52,5	53,5
2013	54,5	55,5	56,5	57,5	58,5	59,5	60,5	61,5	62,5	63,5	64,5	65,5
2014	66,5	67,5	68,5	69,5	70,5	71,5	72,5	73,5	74,5	75,5	76,5	77,5
2015	78,5	79,5	80,5	81,5	82,5	83,5	84,5	85,5	86,5	87,5	88,5	89,5
2016	90,5	91,5	92,5	93,5	94,5	95,5	96,5	97,5	98,5	99,5	100,5	101,5
2017	102,5	103,5	104,5	105,5	106,5	107,5	108,5	109,5	110,5	111,5	112,5	113,5



- 4. Obtención de valores de la Tendencia mensual a través de la recta de regresión estimada:

TENDENCIA MENSUAL	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
1999	33140,7	33212,2	33283,7	33355,2	33426,7	33498,2	33569,7	33641,1	33712,6	33784,1	33855,6	33927,1
2000	33998,6	34070,1	34141,6	34213,1	34284,6	34356,1	34427,6	34499,0	34570,5	34642,0	34713,5	34785,0
2001	34856,5	34928,0	34999,5	35071,0	35142,5	35214,0	35285,5	35356,9	35428,4	35499,9	35571,4	35642,9
2002	35714,4	35785,9	35857,4	35928,9	36000,4	36071,9	36143,4	36214,8	36286,3	36357,8	36429,3	36500,8
2003	36572,3	36643,8	36715,3	36786,8	36858,3	36929,8	37001,3	37072,7	37144,2	37215,7	37287,2	37358,7
2004	37430,2	37501,7	37573,2	37644,7	37716,2	37787,7	37859,2	37930,6	38002,1	38073,6	38145,1	38216,6
2005	38288,1	38359,6	38431,1	38502,6	38574,1	38645,6	38717,1	38788,5	38860,0	38931,5	39003,0	39074,5
2006	39146,0	39217,5	39289,0	39360,5	39432,0	39503,5	39575,0	39646,4	39717,9	39789,4	39860,9	39932,4
2007	40003,9	40075,4	40146,9	40218,4	40289,9	40361,4	40432,9	40504,3	40575,8	40647,3	40718,8	40790,3
2008	40861,8	40933,3	41004,8	41076,3	41147,8	41219,3	41290,8	41362,2	41433,7	41505,2	41576,7	41648,2
2009	41719,7	41791,2	41862,7	41934,2	42005,7	42077,2	42148,7	42220,1	42291,6	42363,1	42434,6	42506,1
2010	42577,6	42649,1	42720,6	42792,1	42863,6	42935,1	43006,6	43078,0	43149,5	43221,0	43292,5	43364,0
2011	43435,5	43507,0	43578,5	43650,0	43721,5	43793,0	43864,5	43935,9	44007,4	44078,9	44150,4	44221,9
2012	44293,4	44364,9	44436,4	44507,9	44579,4	44650,9	44722,4	44793,8	44865,3	44936,8	45008,3	45079,8
2013	45151,3	45222,8	45294,3	45365,8	45437,3	45508,8	45580,3	45651,7	45723,2	45794,7	45866,2	45937,7
2014	46009,2	46080,7	46152,2	46223,7	46295,2	46366,7	46438,2	46509,6	46581,1	46652,6	46724,1	46795,6
2015	46867,1	46938,6	47010,1	47081,6	47153,1	47224,6	47296,1	47367,5	47439,0	47510,5	47582,0	47653,5
2016	47725,0	47796,5	47868,0	47939,5	48011,0	48082,5	48154,0	48225,4	48296,9	48368,4	48439,9	48511,4
2017	48582,9	48654,4	48725,9	48797,4	48868,9	48940,4	49011,9	49083,4	49154,8	49226,3	49297,8	49369,3

- 5. Gráfico comparativo de los valores reales de la serie *Viajeros en establecimientos hoteleros en La Rioja 1999 – 2017* y la Tendencia mensual estimada:





### 3.3. Cálculo y análisis de la variación estacional

Como ya hemos definido previamente la variación estacional responde a aquellas fluctuaciones u oscilaciones que experimenta la serie con periodicidad inferior a un año. En el caso concreto del número de viajeros mensuales en establecimientos hoteleros en La Rioja de 1999 a 2017, podemos apreciar un grado de estacionalidad “estable”, es decir, el comportamiento de la variable a lo largo del año se repite con un grado de semejanza elevado, año tras año. Por tanto, es posible obtener indicadores para cada mes, denominados **Índices de Variación Estacional (IVE)**. Para su obtención, se desarrollará el método de la razón a la tendencia por el método de medias móviles, aislando así la componente estacional.

- 1. División de los valores k – esimales de la serie temporal  $Y_t$  por el valor obtenido en el análisis de la tendencia por medias móviles  $T_t$  (suponiendo que esta contiene a  $C_t$ ), dando lugar a valores sin tendencia ni ciclo:

$$\frac{Y_t}{T_t} = \frac{T_t C_t \cdot S_t \cdot I_t}{T_t C_t} = S_t \cdot I_t$$

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
1999							1,155	1,448	1,217	1,202	0,769	0,691
2000	0,572	0,791	0,958	1,145	1,098	1,040	1,156	1,319	1,197	1,168	0,943	0,728
2001	0,574	0,719	0,916	1,048	1,108	1,087	1,144	1,330	1,213	1,198	0,924	0,657
2002	0,619	0,762	0,971	0,973	1,222	0,978	1,098	1,433	1,163	1,126	0,914	0,712
2003	0,658	0,711	1,040	1,097	1,076	1,051	1,085	1,346	1,193	1,188	0,937	0,676
2004	0,657	0,727	0,905	1,029	1,211	1,064	1,123	1,306	1,183	1,199	0,886	0,700
2005	0,529	0,743	1,032	0,937	1,164	1,096	1,081	1,352	1,273	1,246	0,841	0,669
2006	0,572	0,751	0,904	1,135	1,152	1,024	1,108	1,341	1,164	1,229	0,968	0,731
2007	0,621	0,801	0,878	1,073	1,162	1,052	0,971	1,236	1,249	1,229	0,957	0,729
2008	0,644	0,771	0,944	0,921	1,216	1,052	1,135	1,383	1,332	1,165	0,858	0,657
2009	0,565	0,644	0,851	1,109	1,305	1,057	1,092	1,329	1,271	1,221	0,878	0,647
2010	0,495	0,714	0,895	1,098	1,183	1,061	1,120	1,375	1,298	1,248	0,821	0,708
2011	0,544	0,682	0,794	1,125	1,175	1,104	1,135	1,397	1,334	1,207	0,806	0,761
2012	0,563	0,616	0,889	1,148	1,158	1,061	1,087	1,361	1,306	1,177	0,894	0,670
2013	0,537	0,668	1,015	0,885	1,217	1,091	1,082	1,382	1,328	1,192	0,930	0,697
2014	0,571	0,677	0,834	1,003	1,172	1,093	1,080	1,372	1,373	1,168	0,913	0,703
2015	0,580	0,641	0,906	1,033	1,180	1,078	1,103	1,405	1,342	1,154	0,883	0,736
2016	0,590	0,693	0,912	0,917	1,190	1,040	1,178	1,414	1,278	1,176	0,827	0,782
2017	0,559	0,684	0,825	1,142	1,088	1,063						

- 2. Los valores obtenidos sin tendencia ni ciclo son promediados con valores de cada mes de todos los años, de este modo se considera que la componente accidental queda repartida entre todo el periodo, y la componente estacional queda aislada. Dichos promedios serán los Movimientos Estacionales, que pasados a porcentaje supondrán los definitivos IVE:

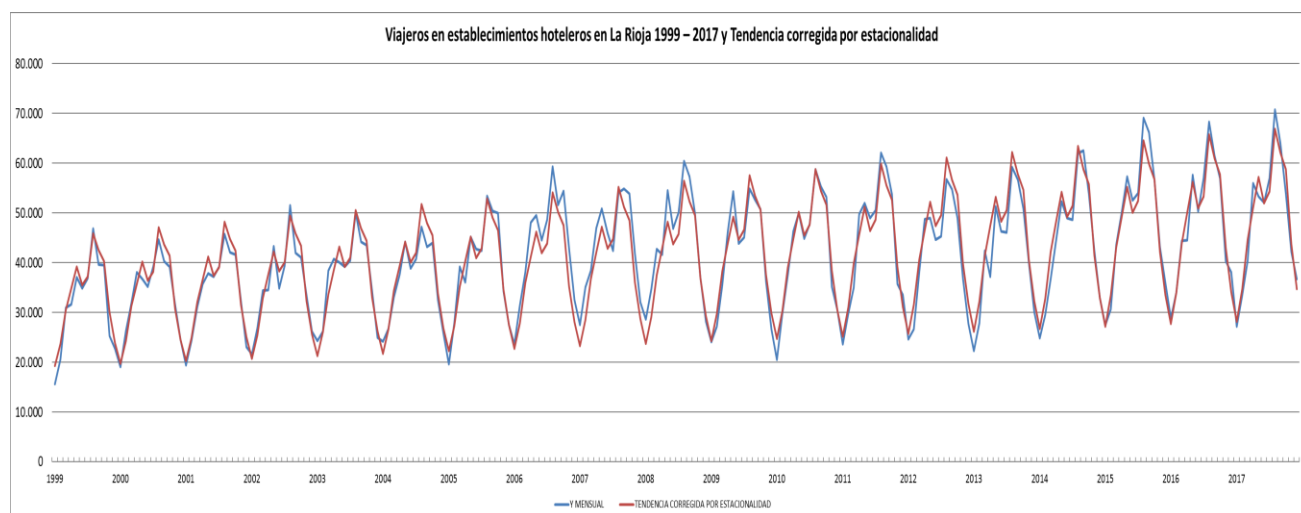
$$IVE(i) = \frac{ME(i)}{\sum_{i=1}^k ME(i)} \cdot k \cdot 100$$

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	TOTAL
ME	0,581	0,711	0,915	1,045	1,171	1,061	1,107	1,363	1,262	1,194	0,886	0,703	11,99852
IVE %	58,1%	71,1%	91,5%	104,6%	117,1%	106,1%	110,7%	136,3%	126,2%	119,4%	88,6%	70,3%	
IVE	0,581	0,711	0,915	1,046	1,171	1,061	1,107	1,363	1,262	1,194	0,886	0,703	12

- 3. Una vez obtenidos los **Índices de Variación Estacional (IVE)** es posible obtener los valores de Tendencia corregidos por estacionalidad, para ello basta con multiplicar los valores de Tendencia  $T_t$  por su IVE correspondiente, por ejemplo los valores de tendencia para todos los periodos de enero, multiplicarán por el IVE de enero 0,581.

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
1999	19243	23612	30455	34876	39145	35537	37177	45850	42550	40342	30005	23852
2000	19741	24222	31240	35773	40150	36447	38127	47019	43632	41366	30765	24455
2001	20239	24832	32025	36670	41155	37357	39077	48189	44715	42391	31525	25058
2002	20738	25441	32810	37567	42160	38267	40027	49358	45798	43415	32286	25661
2003	21236	26051	33595	38464	43164	39177	40977	50527	46881	44439	33046	26264
2004	21734	26661	34380	39361	44169	40087	41927	51696	47964	45464	33806	26867
2005	22232	27271	35165	40258	45174	40997	42877	52866	49046	46488	34567	27470
2006	22730	27881	35950	41155	46178	41907	43827	54035	50129	47513	35327	28073
2007	23228	28491	36735	42052	47183	42817	44777	55204	51212	48537	36087	28677
2008	23726	29101	37520	42949	48188	43727	45727	56373	52295	49562	36848	29280
2009	24224	29711	38305	43846	49192	44638	46677	57543	53377	50586	37608	29883
2010	24723	30321	39090	44743	50197	45548	47628	58712	54460	51610	38368	30486
2011	25221	30931	39875	45640	51202	46458	48578	59881	55543	52635	39129	31089
2012	25719	31541	40660	46537	52206	47368	49528	61050	56626	53659	39889	31692
2013	26217	32150	41445	47434	53211	48278	50478	62220	57709	54684	40649	32295
2014	26715	32760	42230	48331	54216	49188	51428	63389	58791	55708	41410	32898
2015	27213	33370	43015	49228	55220	50098	52378	64558	59874	56733	42170	33502
2016	27711	33980	43800	50125	56225	51008	53328	65727	60957	57757	42930	34105
2017	28210	34590	44585	51022	57230	51918	54278	66897	62040	58781	43691	34708

- 4. Gráfico comparativo de los valores reales de la serie *Viajeros en establecimientos hoteleros en La Rioja 1999 – 2017* y la Tendencia corregida por estacionalidad



### 3.4. Desestacionalización de la serie

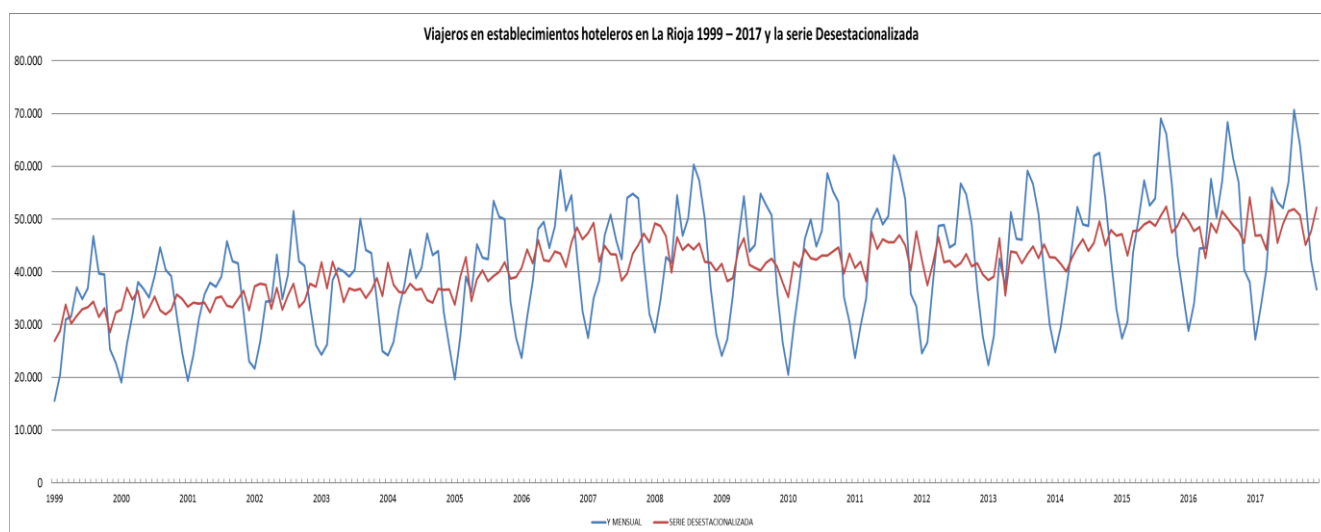
Para eliminar la componente estacional de la serie, recurriremos de nuevo a los IVE, mediante la división de los valores originales entre dichos índices:

$$D_t = \frac{Y_t}{S_t} = \frac{Y_t}{IVE(i)}$$

Los valores obtenidos corresponderán con la serie de datos desestacionalizada  $D_t$ , y consistirán en aquellos valores que se hubieran observado ante la ausencia de influencia estacional.

$D_t$	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE
1999	26860	28807	33805	30201	31614	32866	33302	34334	31433	33103	28545	32352
2000	32770	37012	34744	36356	31366	33137	35358	32738	31952	32828	35674	34872
2001	33347	34158	33932	34149	32357	35003	35338	33571	33295	34844	36405	32716
2002	37276	37698	37588	32951	36963	32849	35538	37774	33252	34420	37767	37152
2003	41786	36868	41958	38923	34221	36894	36446	36745	34989	36478	38816	35430
2004	41695	37567	36190	35994	37752	36583	36811	34633	34169	36791	36620	36667
2005	33742	39074	42781	34490	38579	40258	38263	39213	39977	41819	38675	39050
2006	40753	44229	41636	46003	42239	41994	43890	43471	40914	45610	48445	46198
2007	47354	49311	41932	44940	43408	43303	38311	39661	43438	45132	47270	45579
2008	49183	48716	46723	39854	46564	44138	45241	44284	45434	41882	41749	40135
2009	41509	38223	38834	44133	46371	41355	40719	40237	41735	42449	40958	37903
2010	35269	41818	40964	44245	42622	42290	43081	43086	43828	44577	39707	43415
2011	40720	41884	38197	47548	44379	46191	45636	45557	46983	45015	40356	47667
2012	42320	37466	41601	46578	41775	42086	40945	41624	43371	41022	41603	39457
2013	38383	39130	46393	35536	43832	43631	41634	43400	44842	42618	45178	42809
2014	42673	41569	40107	42559	44626	46154	43934	45499	49531	45043	47892	46839
2015	47134	43057	47708	47860	48973	49519	48724	50665	52409	47445	48687	51112
2016	49634	47713	48534	42542	49214	47421	51474	50129	48745	47694	45498	54123
2017	46867	46965	44192	53516	45468	49076	51459	51898	50707	45105	47617	52186

- Gráfico comparativo de los valores reales de la serie *Viajeros en establecimientos hoteleros en La Rioja 1999 – 2017* y la serie Desestacionalizada:



### 3.5 Predicción del Número de Viajeros en La Rioja. Enero – Mayo 2018

Una vez llevado a cabo el estudio y análisis completo mediante el modelo clásico de la serie *Viajeros en establecimientos hoteleros en La Rioja 1999 – 2017*, y habiendo establecido un modelo de tendencia mensual  $T_t^{(12)} = \frac{495060105}{12} + \frac{10294805}{12^2} \cdot t''$  y los Índices de Variación Estacional para cada mes del año, es posible llevar a cabo uno de los objetivos perseguidos por el análisis como es la predicción de valores futuros para la variable estudiada.

En este caso, tras haber incluido en el modelo todos los valores de las dos últimas dos décadas, 1999 – 2017, en el modelo obtenido, la predicción corresponderá con los valores de los 5 meses sucesivos a la misma, es decir, enero, febrero, marzo, abril y mayo de 2018. El motivo de limitar la predicción a dichos 5 meses es que ya están disponibles los valores reales para estos periodos, y por tanto será posible la comparativa entre la variable real y la estimación de la predicción.

Para llevar a cabo dicha predicción y su cálculo será necesario recurrir al análisis y hacer uso de algunos de los cálculos obtenidos durante el mismo, en primer lugar, la tendencia por el método mínimo cuadrático,

$$T_t^{(12)} = \frac{495060105}{12} + \frac{10294805}{12^2} \cdot t''$$

sustituyendo en ella los valores en números enteros y trasladados por  $t''$ :

Transformación a números enteros					
	ENERO	FEBRERO	MARZO	ABRIL	MAYO
2018	120	121	122	123	124

Traslado = (t- 5,5)					
	ENERO	FEBRERO	MARZO	ABRIL	MAYO
2018	114,5	115,5	116,5	117,5	118,5

Obteniendo los siguientes valores de tendencia:

Tendencia mensual	ENERO	FEBRERO	MARZO	ABRIL	MAYO
2018	49440,8088	49512,3005	49583,7922	49655,2839	49726,7756

Por último, una vez obtenidos los valores de tendencia para los meses indicados, serán multiplicados por sus Índices de Variación Estacional correspondientes,

	ENERO	FEBRERO	MARZO	ABRIL	MAYO
IVE	0,5806484	0,7109347	0,9150235	1,0455808	1,1710855

obteniendo los siguientes valores como predicción:

PREDICCIÓN	ENERO	FEBRERO	MARZO	ABRIL	MAYO
2018	28708	35200	45370	51919	58234

*Comparativa de la predicción obtenida y los valores reales del Número de Viajeros en establecimientos hoteleros en La Rioja. Enero – Mayo 2018.*

Los valores reales del Número de Viajeros en establecimientos hoteleros en La Rioja para los 5 primeros meses de 2018 son los siguientes:

Nº Viajeros La Rioja	ENERO	FEBRERO	MARZO	ABRIL	MAYO
2018	26708	31371	41839	47645	53411

La comparación de los valores obtenidos por la predicción y los valores reales de la serie evidencian una más que aceptable aproximación entre ambos valores. En líneas generales puede establecerse y apreciarse de un modo claro que los valores de la predicción son aproximadamente un 9% mayores que los valores reales de la variable. Esta dispersión podemos explicarla como la componente aleatoria presente en cualquier modelo de serie temporal.

Error Predicción	ENERO	FEBRERO	MARZO	ABRIL	MAYO	Error medio
2018	7,487%	12,206%	8,440%	8,970%	9,031%	<b>9,2267%</b>

Además, cabe recordar el cálculo e interpretación del Coeficiente de correlación  $R^2 = 0.861$ , según el cual establecíamos la obtención de un coeficiente muy próximo a 1, y por tanto que el modelo de tendencia calculado era muy representativo y de un grado de calidad elevado. Dicho modelo de tendencia además es corregido por estacionalidad gracias a la utilización de los Índices de Variación Estacional mensuales, cuyo valor representativo recae en la base de datos utilizada de prácticamente dos décadas, cuyo valor es fiable y permite así obtener aproximaciones cercanas a los valores reales de la variable mediante una predicción con errores de entorno al 9%.

## 4. CONCLUSIONES

Tras la realización y exposición del análisis estadístico de series temporales llevado a cabo sobre un caso real de la economía riojana, concretamente sobre el Número de Viajeros en La Rioja entre 1999 – 2017, podemos alcanzar una serie de conclusiones que aporten sentido a los resultados obtenidos a lo largo del estudio.

En primer lugar, resaltar la validez de la Teoría Clásica de Series Temporales, pese a su sencillez a la hora de simplificar el modelo para una variable concreta en un horizonte temporal determinado, permite analizar y estudiar de un modo más que válido el comportamiento de dicha variable y su asociación con el transcurso de la unidad temporal en que se presenta. Desde un punto de vista crítico, cabe decir que se trata de una Teoría prácticamente obsoleta, aunque se posiciona como la base de cualquier modelo de series temporales más actualizado.

En cuanto al caso de estudio desarrollado sobre una variable propia del sector turístico riojano en este caso, debemos obtener conclusiones que traten de explicar el comportamiento de este sector durante las dos últimas décadas en nuestra comunidad autónoma. Las principales características del sector y la variable analizada podemos asociarlas a dos ideas principales, **crecimiento y estacionalidad**.

### *Crecimiento*

En primer lugar y como conclusión principal, el turismo riojano está comportándose de un modo positivo, siguiendo una línea creciente que parece haberse asentado durante las dos últimas décadas. Para tratar de refrendar dicha afirmación basta con hacer referencia a la Tendencia anual obtenida durante el análisis.

$$T_t = 495060,105 + 10294,805 \cdot t'$$

La propia obtención de la tendencia anual ( $T_t$ ) permite asegurar que la variable estudiada experimenta un crecimiento continuado de más de 10.000 viajeros al año de media, dato que puede traducirse en la consolidación del sector en nuestra región y su crecimiento progresivo en volumen de viajeros.

### *Estacionalidad*

Otro de los aspectos clave que se obtiene del análisis elaborado es la fuerte y más que definida estacionalidad del sector y lo acentuado que queda el comportamiento de la variable dentro del año natural. La Rioja, pese a no caracterizarse por ser una región propiamente turística en nuestro país, ha sido capaz de reinventarse y sacarle el máximo provecho a las posibilidades turísticas con las que cuenta, turismo cultural, turismo de naturaleza o turismo gastronómico, principalmente. Sin embargo, y aunque parezca que la estacionalidad en estas modalidades de turismo no viene tan

definida, es un hecho que la estacionalidad ha venido acentuándose a lo largo de las últimas décadas, como puede apreciarse en la mayor disparidad entre máximos y mínimos anuales. Es uno de los motivos que, a su vez, genera que el modelo que mejor explica el comportamiento de la variable es el multiplicativo ya que propicia una mayor amplitud y dispersión en cuanto a la tendencia, conforme la variable avanza a lo largo de la unidad temporal.

Los resultados que apoyan y refrendan la estacionalidad como uno de los aspectos clave en la explicación del modelo recaen sobre la obtención llevada a cabo de los Índices de Variación Estacional (IVE), los cuales nos permiten dividir el comportamiento de la variable, en este caso en meses, ya que la variable se comporta de un modo muy semejante para cada mes del año, a lo largo de los 19 años estudiados, definiéndose así la estacionalidad del modelo. Los valores obtenidos de los mismos deben interpretarse como la dispersión existente entre el valor de la Tendencia, un valor que suaviza el comportamiento de la variable y que no tiene en cuenta la estacionalidad de la misma, y el valor que toma la variable en función del mes en el que nos encontremos,

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SEPTIEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	TOTAL
ME	0,581	0,711	0,915	1,045	1,171	1,061	1,107	1,363	1,262	1,194	0,886	0,703	11,99852
IVE %	58,1%	71,1%	91,5%	104,6%	117,1%	106,1%	110,7%	136,3%	126,2%	119,4%	88,6%	70,3%	
IVE	0,581	0,711	0,915	1,046	1,171	1,061	1,107	1,363	1,262	1,194	0,886	0,703	12

Los valores obtenidos para los IVE, permiten apreciar la marcada estacionalidad y la dispersión entre meses, ya que existe una diferencia de más de un 70% respecto a la tendencia entre el mes de menor recepción media de viajeros, enero que se sitúa más de un 40% por debajo de la Tendencia, y agosto que se posiciona con más de un 35% de viajeros de media por encima de dicho parámetro.

### *Predicción*

Por último, cabe destacar la validez y calidad del modelo obtenido, cuya finalidad principal recae en la posibilidad de ofrecer predicciones para un futuro cercano. La predicción llevada a cabo puede interpretarse como más que válida y de un buen grado de aproximación al comportamiento real de la variable.

$$T_t^{(12)} = \frac{495060105}{12} + \frac{10294805}{12^2} \cdot t''$$

Los valores que arroja la predicción llevada a cabo para los 5 primeros meses de 2018 muestran un error en torno al 10%, cifra que podemos justificar como la variación accidental existente en cualquier modelo de serie temporal, y que como su propio nombre indica no es posible medir ni anticipar en el proceso predictivo. De este modo, y en base a la predicción y comparación con los valores reales de la variable, puede afirmarse que el Número de Viajeros en La Rioja continuará con el comportamiento de crecimiento y de estacionalidad explicado por el modelo.

## 5. BIBLIOGRAFÍA

JESÚS ESTEBAN GARCÍA. (2005): “Estadística descriptiva y nociones de probabilidad”. Editorial: Thomson.

VENANCIO TOMEIO PERUCHA e ISAÍAS UÑA JUÁREZ. (2003): “Lecciones de estadística descriptiva: curso teórico – práctico”. Editorial: Thompson.

WEB GOBIERNO DE LA RIOJA: <https://www.larioja.org/es>

- Instituto de estadística de La Rioja:

<http://www.larioja.org/estadistica/es>

- LA RIOJA TURISMO: <https://lariojaturismo.com/>

WEB INSTITUTO NACIONAL DE ESTADÍSTICA: <http://www.ine.es/>

- Encuesta de ocupación hotelera:

[http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica\\_C&cid=1254736177015&menu=ultiDatos&idp=1254735576863](http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736177015&menu=ultiDatos&idp=1254735576863)

- Base de datos:

<http://www.ine.es/jaxiT3/Tabla.htm?t=2074&L=0>



